Curvelet-based non-linear adaptive subtraction with sparseness constraints

Felix J Herrmann, Peyman P Moghaddam
(EOS-UBC)
Spiky/minimal structure deconvolution (Claerbout, Ulrych, Oldenburg, Sacchi, Trad etc.)

Sparse Radon (Ulrych, Sacchi, Trad, etc.)

FFT/Focus transform-based Interpolation (Duyndam, Zwartjes, Verschuur, Berkhout)

Redundant dictionaries/Morphological component separation/Pursuits (Mallat, Chen, Donoho, Starck, Elad)

L2-migration (Nemeth, Chavent, de Hoop, Hu, Kuehl)

2-D/3-D Curvelets Non-linear synthesis (Durand, Starck, Candes, Demanet, Ying)

Anisotropic Diffusion (Osher)
Goals

Processing & imaging scheme

⭐ increases resolution & SNR
⭐ preserves edges = freq. content
⭐ works with and extends existing

• noise removal/signal separation
• imaging schemes

Develop the right *language* to deal with SNR ≤ 0 ....
General Framework

Divide-and-conquer approach:

1. Sparseness with thresholding
2. Continuity with constrained optimization

Main focus:

- use of Curvelets as optimal Frames
- (iterative) thresholding
Appetizer

missing data with multiples  estimated primaries  estimated multiples
Wish list

Seek a *transform domain* that is

- relative insensitive to *local phase*
- sparse & *local* (position/dip)
- *optimal* for *curved* events
- *well-behaved* under *operators*
- *near diagonalizes* *Covariance*

Aim to bring out those high frequencies with *ultra-low* SNR<0!
Why curvelets

- Nonseparable
- Local in 2-D space
- Local in 2-D Fourier
- Anisotropic
- Multiscale
- Almost orthogonal
- Tight frame $B^TB = I$
- Optimal
3-D Curvelets

Thanks to Demanet & Ying
3-D Curvelets

Thanks to Demanet & Ying
3-D Curvelets

Thanks to Demanet & Ying
Why curvelets

\[ W_j = \{ \xi, \ 2^j \leq |\xi| \leq 2^{j+1}, |\theta - \theta_j| \leq \pi \cdot 2^\lfloor j/2 \rfloor \} \]

Fourier/SVD/KL
\[ \| f - \tilde{f}_m^F \| \propto m^{-1/2}, \ m \to \infty \]

Wavelet
\[ \| f - \tilde{f}_m^W \| \propto m^{-1}, \ m \to \infty \]

Optimal data adaptive
\[ \| f - \tilde{f}_m^A \| \propto m^{-2}, \ m \to \infty \]

Close to optimal Curvelet
\[ \| f - \tilde{f}_m^C \| \leq C \cdot m^{-2} (\log m)^3, \ m \to \infty \]

source: Candes'01, Stein '90
Why curvelets

Curvelet in FK-domain
Denoising

Input data with noise

Curvelet transform

Threshold

Inv. curvelet transform

Consequence of curvelet transform

$S_\lambda(Bd)$

Bd

denoised Curvelet coeff.

Filtered input data
Denoising
Denoising
Wavelets
Curvelets
Coherent noise suppression with curvelets

Input data with noise

Threshold

Curvelet transform

Filtered input data

Curvelet coeff. pred. noise

Inv. curvelet transform

Curvelet coeff. model

\[ S_{\lambda}(Bd) \]

\[ |B\hat{n}| \]
Multiple suppression with curvelets

Input with multiples
Multiple suppression with curvelets

Output curvelet filtering with stronger threshold

Preserved primaries
Denoising

Denoising $\Leftrightarrow$ signal separation

- Multiple & Ground-roll removal
- Migration denoising (H & M ‘04)
- 4-D difference cubes

Model “noise”

Strategies:

- Weighted thresholding
- Iterated weighted thresholding
L2/L1-matched filter

Matched filter:
\[ \hat{n} : \min_{\Phi} = \| d_{\text{noisy data}} - \Phi t \|^p \times m_{\text{pred. noise}} \]

- \( p=1 \) enhances sparseness
- residue is the denoised data
- risk of over fitting

Loose primary reflection events ...
Colored denoising

\[\hat{m} = \arg\min_m \frac{1}{2} \left\| C_n^{-1/2} (d - m) \right\|_2^2 + J(m)\]

with covariance

\[C_n \equiv \mathbb{E}\{nn^T\}\]

and both \(m, n\) related to PDE
Ground-roll removal with curvelets

Radon

Iterations=3
Weighted thresholding

Covariance model & noise near diagonal:

$$BC_n \text{ or } mB^T \approx \Gamma^2 \quad \text{near diagonal}$$

For ortho basis and app. noise prediction:

$$\hat{m} = B^T S_\lambda \Gamma (Bd)$$

equivalent to

$$\hat{m} = B^T \arg\min_{\tilde{m}} \frac{1}{2} \left\| \Gamma^{-1} (\tilde{d} - \tilde{m}) \right\|_2^2 + \left\| \tilde{m} \right\|_{1,\lambda}$$

$$\tilde{d} = Bd, \quad \tilde{m} = Bm \quad \text{and} \quad \Gamma = \left[ \text{diag}\{\text{diag}\{B\hat{n}\}\} \right]^{1/2}$$
Multiple suppression with curvelets

Output curvelet filtering with stronger threshold

Preserved primaries
4-D difference
3-D Curvelets
Iterative thresholding

Curvelets are Frames:

- redundant (factor 7.5-4)
- thresholding does not solve:

\[
\hat{m} = B^T \arg \min_{\tilde{m}} \frac{1}{2} \| \Gamma^{-1} (\tilde{d} - \tilde{m}) \|_2^2 + \| \tilde{m} \|_{1,\lambda}
\]

Alternative formulation by iterative thresholding!
Iterative thresholding

Solves signal-separation problem:

\[ s = s_1 + s_2 + n \quad \text{and} \quad s = s_1 + s_2 \]

by minimizing LP-program

\[ \hat{x} = \arg \min_x \| s - \Phi x \|^2_2 + \| x_1 \|_{w_1,1} + \| x_2 \|_{w_2,1} \]

with

\[ x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T \quad \text{and} \quad \Phi = \begin{bmatrix} \Phi_1 & \Phi_2 \end{bmatrix} \]

Starck et al.; Daubechies et al., Donoho 2004
(also Zwartjes, Sacchi, Trad & Ulrich)
Example shallow water environment

Input data with multiples

Result L.S. subtraction

Leakage of L.S. subtraction
Example shallow water environment

Input data with multiples

Result curvelet-based subtraction
Example shallow water environment

Input data with multiples

True primaries
Example shallow water environment

Input data with multiples and noise

Result curvelet-based subtraction
Global optimization

After thresholding:
- remove artifacts & ‘miss fires’
- normal operator (inversion)

Impose additional *penalty functional*
- *prior* information
- *sparseness & continuity*

*Defining the right norm is crucial ...*
Global optimization

Formulate constrained optimization:

\[ \hat{m} : \min_m J(m) \quad \text{s.t.} \quad |\tilde{m} - \hat{m}_0|_\mu \leq e_\mu, \quad \forall \mu \]

with

\[ \hat{m}_0 = B^\dagger \Theta \lambda \Gamma (\tilde{d}) \]

and with \( e_\mu \) threshold and noise-dependent *tolerance* on curvelet coeff.
Imaging with Curvelets

Least-squares migrated Image

Constrained Optimization
Penalty functionals

Anisotropic diffusion for imaging:

\[ J(m) = \| \Lambda^{1/2} \nabla m \|_p \]

with

\[ \Lambda[\bar{c}] = \frac{1}{|\nabla \bar{c}|^2 + 2\beta^2} \left\{ \begin{pmatrix} \partial_{x_2} \bar{c} \\ -\partial_{x_1} \bar{c} \end{pmatrix} \begin{pmatrix} \partial_{x_2} \bar{c} & -\partial_{x_1} \bar{c} \end{pmatrix} + \beta^2 I \right\} \]

brings out wavefront set.
Denoising
Denoising
Denoising
Presented a framework that

**Stable signal separation via thresholding:**
- Ground-roll & Multiple removal (Wednesday)
- Compute 4-D difference Cubes

**Improves imaging & inversion:**
- deals with incoherent noise & missing data
- sparseness constrained imaging (H & P ‘04)
- extended to 3-D
Acknowledgements

Candes, Donoho, Demanet, Ying for making their Curvelet code available.

Partially supported by a NSERC Grant.